Supporting Information for

Ideas and perspectives: Beyond Microbes: Integrating Termites into Global Soil Carbon Cycling Models

Umar Farooq[1], Chiara Pasut[2], Ying-Ping Wang[3], Amy E. Zanne[4], Habacuc Flores-Moreno[5], Baptiste Joseph Wijas[4,6], David I. Forrester[7], Jacqueline R. England[3], Bennett Macdonald[1], Zachary A. Brown[1], Senani Karunaratne[1]

[1] CSIRO Agriculture and Food, Canberra, Australian Capital Territory, Australia,
[2] CSIRO Agriculture and Food, Waite Campus, South Australia, Australia,
[3] CSIRO Environment, Clayton South, Victoria, Australia,
[4] Cary Institute of Ecosystem Studies, Millbrook, New York, USA,
[5] CSIRO Health and Biosecurity, Dutton Park, Queensland, Australia,
[6] School of Environment, University of Queensland, Brisbane, Queensland, Australia,
[7] CSIRO Environment, Canberra, Australian Capital Territory, Australia,

Contents of this file

Text S1. Model Formulation

Litter supply (I) was calculated from net primary production (NPP) using vegetation-specific allocation fractions and component-specific residence times(Wang et al., 2010). Litter production (g C m^{-2} d^{-1}) was partitioned into I_M, I_S, and I_{CWD} as follow:

$$I \rightarrow I_M = \frac{0.4\,\alpha_{\mathrm{leaf}}}{\tau_{\mathrm{leaf}}}\mathrm{NPP} + \frac{0.3\,\alpha_{\mathrm{root}}}{\tau_{\mathrm{root}}}\mathrm{NPP}, \quad \text{S1}$$

$$I \rightarrow I_S = \frac{0.6\,\alpha_{\mathrm{leaf}}}{\tau_{\mathrm{leaf}}}\mathrm{NPP} + \frac{0.7\,\alpha_{\mathrm{root}}}{\tau_{\mathrm{root}}}\mathrm{NPP} + \frac{0.2\,\alpha_{\mathrm{wood}}}{\tau_{\mathrm{wood}}}\mathrm{NPP}, \quad \text{S2}$$

$$I \rightarrow I_{\mathrm{CWD}} = \frac{0.8\,\alpha_{\mathrm{wood}}}{\tau_{\mathrm{wood}}}\mathrm{NPP}, \quad \text{S3}$$

where α_{leaf}, α_{root}, and α_{wood} are vegetation-specific NPP allocation coefficients for leaf, root, and woody litter, respectively, and τ_{leaf}, τ_{root}, and τ_{wood} are the corresponding litter residence times. Allocation coefficients, residence times, and steady state litter inputs to I_M, I_S, and I_{CWD} are summarized in Table S1. Each litter pool underwentmicrobial decomposition using first-order kinetics modulated by temperature and soil moisture:

$$I_M \rightarrow \mathrm{MIC}_{R,k} = k_M\, f(T)\, f(\theta)\, I_M, \quad \text{S4}$$

$$I_S \rightarrow \mathrm{MIC}_{R,k} = k_S\, f(T)\, f(\theta)\, I_S, \quad \text{S5}$$

$$I_{\mathrm{CWD}} \rightarrow \mathrm{MIC}_{R,k} = k_{\mathrm{CWD}}\, f(T)\, f(\theta)\, I_{\mathrm{CWD}}, \quad \text{S6}$$

where k_M, k_S, and k_{CWD} represent first-order decay constants for I_M, I_S, and I_{CWD} pools, respectively (Table S2). Temperature sensitivity followed a standard Q_{10} formulation:

$$f(T) = Q_{10}^{\frac{T-T_{\mathrm{ref}}}{10}}, \quad \text{A}$$

where Q_{10} is a dimensionless sensitivity parameter, T_{ref} (°C) is the reference temperature, and T (°C) is the ambient temperature (°C) (Table S2). Moisture limitation was represented using a piecewise scalar with minimum (θ_{min}), optimal (θ_{opt}), and maximum (θ_{max}) soil moisture thresholds:

$$f(\theta) = \begin{cases} 0, & \theta \le \theta_{\mathrm{min}} \\ \frac{\theta - \theta_{\mathrm{min}}}{\theta_{\mathrm{opt}} - \theta_{\mathrm{min}}}, & \theta_{\mathrm{min}} \le \theta \le \theta_{\mathrm{opt}} \\ 1, & \theta_{\mathrm{opt}} \le \theta \le \theta_{\mathrm{max}} \end{cases}. \quad \text{B}$$

Termites-driven litter consumption fluxes were given by:

$$I_{\mathrm{CWD}} \rightarrow \mathrm{TER}_{\mathrm{XP}} = k_{\mathrm{ing},i}\, \mathrm{TER}_{\mathrm{XP}}\,, \quad \text{S7}$$

$$I_S \rightarrow \mathrm{TER}_{\mathrm{FG}} = k_{\mathrm{ing},i}\, \mathrm{TER}_{\mathrm{FG}}\,, \quad \text{S8}$$

$$\mathrm{LMWC} \rightarrow \mathrm{TER}_{\mathrm{SF}} = k_{\mathrm{ing},i}\, \mathrm{TER}_{\mathrm{SF}}\,, \quad \text{S9}$$

where $k_{\mathrm{ing},i}$ (g C g^{-1} termite C d^{-1}) is the per-biomass consumption rate for feeding guild i, and $\mathrm{TER}_{\mathrm{XP}}$, $\mathrm{TER}_{\mathrm{FG}}$, and $\mathrm{TER}_{\mathrm{SF}}$ (g C m^{-2}) are the biomass densities of the respective termite guilds. The flux of termite-driven CH_4 to MIC_m is:

$$\begin{aligned} F_{\mathrm{CH_4}} \rightarrow \mathrm{MIC}_m &= \varepsilon_{\mathrm{meth}}\, f_{\mathrm{ox}}\, F_{\mathrm{CH_4}}, \\ F_{\mathrm{CH_4 \rightarrow CO_2}} &= (1 - \varepsilon_{\mathrm{meth}}) f_{\mathrm{ox}}\, F_{\mathrm{CH_4}} \end{aligned} \quad \text{S10}$$

where $F_{\mathrm{CH_4}}$ (g C d^{-1}) is the gross CH_4 produced by termites, f_{ox} is the oxidation fraction, and $\varepsilon_{\mathrm{meth}}$ is the methanotrophic carbon use efficiency. Necromass production from each feeding guild was represented as:

$$\mathrm{TER}_i \rightarrow \mathrm{TER}_n = \frac{\mathrm{TER}_i}{\tau_{B,i}}, \quad \text{S11}$$

where TER_i is the biomass of guild i and $\tau_{B,i}$ (d) is the corresponding biomass turnover time. Carbon fluxes from frass and necromass pools were partitioned between labile and stabilized fractions as follows:

$$\begin{aligned} \mathrm{TER}_f \rightarrow \mathrm{LMWC} &= \omega\ k_f\ f(T)\, f(\theta)\, \mathrm{TER}_f\,, \\ \mathrm{TER}_f \rightarrow \mathrm{MAOC} &= (1-\omega)\ k_f\ f(T)\, f(\theta)\, \mathrm{TER}_f\,, \end{aligned} \quad \text{S12}$$

$$\begin{aligned} \mathrm{TER}_n \rightarrow \mathrm{LMWC} &= \delta\ k_B\ f(T)\, f(\theta)\, \mathrm{TER}_n\,, \\ \mathrm{TER}_n \rightarrow \mathrm{MAOC} &= (1-\delta)\ k_B\ f(T)\, f(\theta)\, \mathrm{TER}_n\,, \end{aligned} \quad \text{S13}$$

where ω and δ determine the fractions of decomposed carbon entering the LMWC and MAOC pools, respectively, k_f and k_n are the first-order turnover rate constants for frass and necromass, modulated by temperature and moisture.

Table S1. Ecosystem-specific NPP allocation coefficients and plant tissue mean residence times.

Vegetation type	NPP allocation coefficients			Mean residence time (yr)			Steady State litter inputs		
	$\alpha_{\mathbf{leaf}}$	$\alpha_{\mathbf{root}}$	$\alpha_{\mathbf{wood}}$	$\tau_{\mathbf{leaf}}$	$\tau_{\mathbf{root}}$	$\tau_{\mathbf{wood}}$	I_M (Pg C yr^{-1})	I_S (Pg C yr^{-1})	$I_{\mathbf{CWD}}$ (Pg C yr^{-1})
Tropical evergreen	0.34	0.45	0.21	1.75	14	65	6.26	12.96	3.88
Tropical deciduous	0.35	0.40	0.25	0.80	10	60	1.41	2.93	1.09
Temperate broadleaf	0.28	0.57	0.15	1.10	10	50	0.36	0.76	0.15
Temperate needleleaf	0.41	0.27	0.32	1.40	14	75	0.46	0.95	0.49
Temperate deciduous	0.35	0.40	0.25	0.80	10	60	0.79	1.65	0.61
Mixed forest	0.35	0.25	0.40	1.20	10	50	0.50	1.08	0.75
Savanna	0.25	0.65	0.10	1.50	4	40	3.45	7.30	0.93
Grassland	0.30	0.70	–	1.0	3	0	1.02	2.06	0.00
Dense shrubland	0.40	0.45	0.15	1.00	5	40	0.81	1.62	0.33
Open shrubland	0.40	0.45	0.15	1.00	5	40	0.43	0.85	0.17
Total							15.50	32.16	8.40

Table S2. Model Parameters.

Parameter	Description	Plausible range
k_M	Metabolic litter turnover rate constant	5.0 – 7.0 yr^{-1} (Parton et al., 1987)
k_S	Structural litter turnover rate constant	1.7 – 2.3 yr^{-1} (Parton et al., 1987)
$k_{\mathbf{CWD}}$	CWD turnover rate constant	0.08 – 0.12 yr^{-1} (Zell et al., 2009)
Q_{10}	Dimensionless temperature sensitivity constant	2
$T_{\mathbf{ref}}$	Reference temperature for temperature scalar	10 °C
$\theta_{\mathbf{min}}$	Minimum moisture content threshold	0.05
$\theta_{\mathbf{opt}}$	Optimum moisture content threshold	0.6
$\theta_{\mathbf{max}}$	Maximum moisture content threshold	0.9
$\varepsilon_{\mathrm{meth}}$	Methanotrophs carbon use efficiency	0.4 – 0.6

Table S3. Vegetation-specific parameters from (Sanderson, 1996) used to estimate termite biomass and its partitioning among feeding guilds. Here, λ_{SF}, λ_{FG}, and λ_{XP} represent the fraction of total termite biomass assigned to TER_{SF}, TER_{FG}, and TER_{XP} termites, respectively.

Vegetation type	γ (g m^{-2})	b (g^{-1} C m^2 d^1)	Asia & Africa			America & Australia		
			λ_{SF}	λ_{FG}	λ_{XP}	λ_{SF}	λ_{FG}	λ_{XP}
Tropical evergreen	1.21	0.29	0.31	0.10	0.50	0.31	–	0.69
Tropical deciduous	1.21	0.29	0.31	0.10	0.50	0.31	–	0.69
Temperate broadleaf	0.70	0.40	–	–	1.0	–	–	1.0
Temperate needleleaf	1.00	0.44	–	–	1.0	–	–	1.0
Temperate deciduous	0.80	0.47	–	–	1.0	–	–	1.0
Mixed forest	1.10	0.44	–	–	1.0	–	–	1.0
Savanna	1.10	0.40	–	–	0.33	0.07	–	0.93
Grassland	1.21	0.47	1.0	–	0	1.0	–	–
Dense shrubland	1.21	0.51	–	–	0.33	0.07	–	0.93
Open shrubland	0.50	0.36	–	0.66	0.34	0.07	–	0.93

Text S2. Simulation protocol and global sensitivity analysis

Daily forcing data were assembled from observational products spanning 2005–2024. Gross primary productivity (GPP) and net ecosystem productivity (NPP) were obtained from the MODIS 8-day productivity product (Steve and Zhao, 2021), while daily air temperature was obtained from ERA5 reanalysis (ERA5, 2017), and soil moisture was obtained from the SMAP dataset for 2016–2023 (ONeill et al., 2021). These datasets were used to construct daily climatological forcing time series, which are repeated annually during model spin-up.

For each grid cell, the termite module was integrated forward in time under repeated climatological forcing to allow transient pools to approach quasi-steady state. Each simulation used a 1000-year spin-up, allowing litter pools, termite biomass, frass, necromass, and methanotrophic pools to stabilize. Model diagnostics were extracted from the final simulation year. The model produced four primary annual outputs describing termite-mediated carbon fluxes, total CO_2 production, net CH_4 emission, carbon transferred to low-molecular-weight substrates (LMWC), and carbon transferred to mineral-associated organic carbon (MAOC). These outputs formed the basis of the sensitivity analysis.

To quantify parameter importance and interaction effects across the full parameter space, we implemented a two-stage global sensitivity analysis using Morris screening (Morris, 1991) followed by Sobol variance decomposition. The Morris elementary-effects method was first applied to screen all the parameters governing termite biomass dynamics, ingestion rates, carbon partitioning, residue decomposition, and stabilization fractions (Tables 1 and S3). Parameter ranges were defined relative to baseline values reported in Table 1. To reflect varying levels of empirical uncertainty, perturbation ranges were defined as ±15% for productivity–biomass scaling parameters and litter turnover rates, ±30% for ingestion rates and carbon partitioning fractions, and ±50% for methane oxidation and turnover parameters. For parameters lacking direct empirical constraints (e.g., ω, δ , and $\boldsymbol{\tau_B}$), plausible ranges were defined based on a combination of literature-informed bounds from analogous soil and microbial processes, consistency with mass balance constraints, and the need to capture realistic variability without introducing numerical instability. In all cases, ranges were centered on baseline values and selected to span uncertainty levels comparable to those applied in the global sensitivity analysis.

Morris experiments were conducted independently at grid cells selected via stratified random sampling within vegetation classes (up to 500 locations per class). At each location, the Morris design employed 10 random trajectories with 8 discrete levels, resulting in $10(j+1)$ model evaluations per site for a problem with j uncertain parameters. Sensitivity was evaluated for four annual outputs: total CO_2 production, net CH_4 emission, carbon transferred to LMWC, and carbon transferred to MAOC. Site-level indices were aggregated across grid cells to derive vegetation-level mean and variance estimates.

Parameters consistently identified as influential in the Morris screening ($m = 5$) were retained for variance-based Sobol sensitivity analysis (Sobol, 2001). Sobol sampling was performed independently at each selected grid cell using a base sample size of $n = 512$. Using the Saltelli sampling scheme for estimation of first-order and total-order indices, this design required $(m+2)n$ model evaluations per site. First-order Sobol indices (S_1) quantify the fraction of output variance attributable to variation in a single parameter alone, while total-order indices (S_T) capture the total contribution of each parameter including all interaction effects. The

difference $S_T - S_1$ therefore provides a measure of interaction strength. Sobol indices were aggregated across spatial locations to derive vegetation-level sensitivity patterns.

Text S3. Analytical propagation of parameter uncertainty

To quantify how parameter uncertainty propagates through the termite-driven carbon model, we developed an analytical framework based on first-order error propagation. The model structure allowed all fluxes and pools to be expressed as linear functions of a common carbon ingestion flux. Our approach proceeded in three steps: (a) reformulation of the model to isolate independent parameters while enforcing mass balance; (b) expression of all outputs as functions of uncertain parameters; and (c) propagation of parameter uncertainty using analytical derivatives. Parameter ranges were defined based on the bounds defined in Table 1 and were assumed to follow independent uniform distributions. For a parameter x defined over the interval $[L, U]$, the standard deviation is calculated as

$$\sigma_x = \frac{U-L}{\sqrt{12}}, \quad \text{S14}$$

Carbon partitioning fractions were constrained by mass balance. To enforce this constraint while retaining analytical tractability, the fraction corresponding to the analysed output was eliminated. For example, when analysing CO_2 emissions,

$$f_{\mathrm{CO_2}} = 1 - f_B - f_{\mathrm{frass}} - f_{\mathrm{CH_4}}. \quad \text{S15}$$

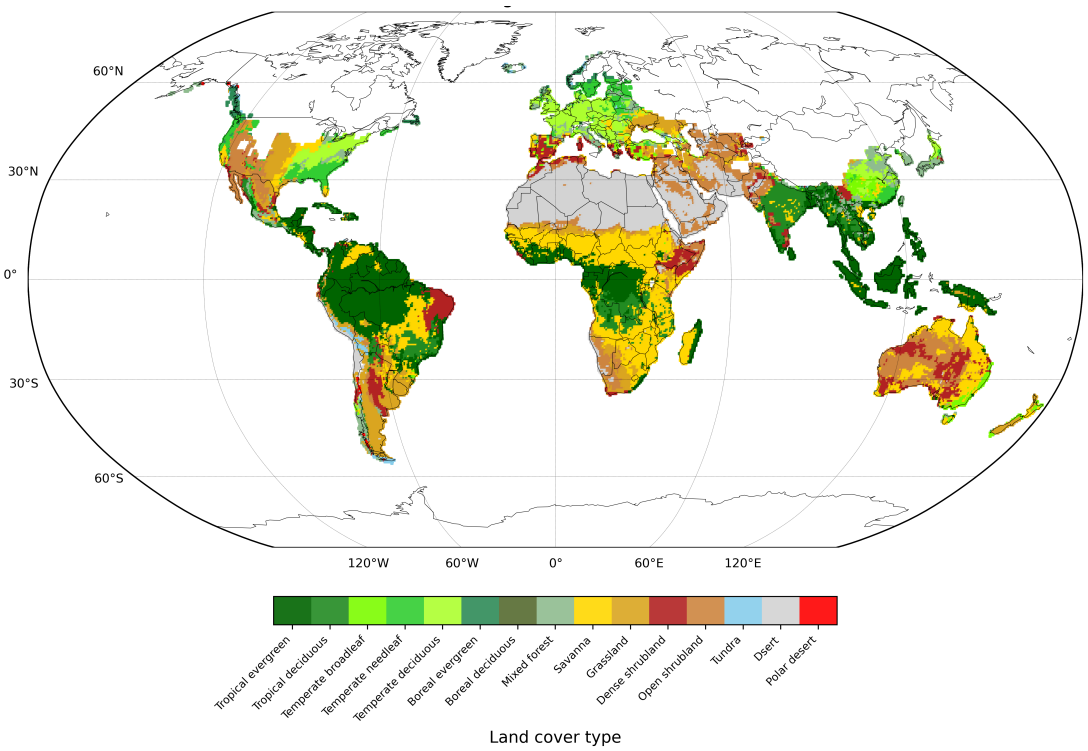
After applying this constraint, the remaining partitioning parameters were treated as independent uncertain variables. All model outputs can be written in the general form

$$Y = k_{\mathrm{ing}}\, K\, g(\theta), \quad \text{S16}$$

where K is a constant environmental scaling factor and $g(\theta)$ is a dimensionless function of partitioning and decay parameters. Because this formulation is linear in all uncertain parameters, uncertainty propagation can be performed analytically using first-order error propagation. Output variance is given by

$$\mathrm{Var}(Y) = \sum_i \left(\frac{\partial Y}{\partial x_i}\right)^2 \sigma_i^2, \qquad \text{S17}$$

where x_i are uncertain parameters and σ_i^2 their variances. Because k_{ing} enters all outputs multiplicatively, its uncertainty contributes proportionally to the variance of all fluxes and pools, whereas other parameters contribute through their influence on the dimensionless function $g(\theta)$.



125 **Figure S1.** Spatial distribution of termite habitat identified using Ramankutty and Foley (1999) global vegetation dataset, excluding regions with monthly minimum air temperatures below −8 °C.

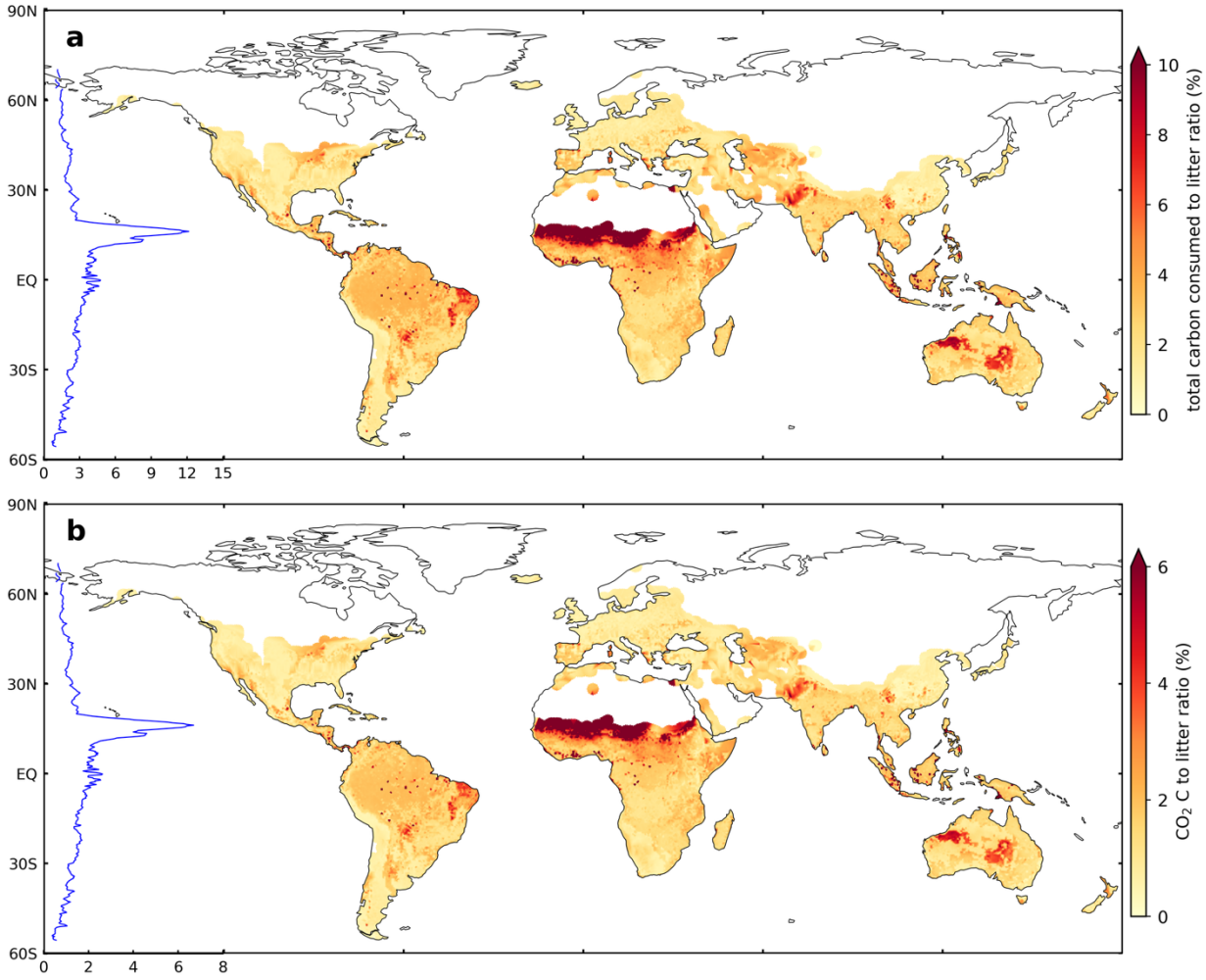


Figure S2. Spatial distribution of (a) the ratio of total carbon consumed by termites to litter input and (b) the ratio of carbon released as CO_2 to litter input. The subset in each plot shows the latitudinal mean of the corresponding variable, using the same units as the associated colorbar.

References

Copernicus Climate Change Service (C3S) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS).

Morris, M.D., 1991. Factorial Sampling Plans for Preliminary Computational Experiments. Technometrics 33, 161–174. https://doi.org/10.1080/00401706.1991.10484804

ONeill, P., Chan, S., Njoku, E., Jackson, T., Bindlish, R., Chaubell, J., Colliander, A., 2021. SMAP Enhanced L3 Radiometer Global and Polar Grid Daily 9 km EASE-Grid Soil Moisture, Version 5. https://doi.org/10.5067/4DQ54OUIJ9DL

Parton, W.J., Schimel, D.S., Cole, C.V., Ojima, D.S., 1987. Analysis of Factors Controlling Soil Organic Matter Levels in Great Plains Grasslands. Soil Sci. Soc. Am. J. 51, 1173–1179. https://doi.org/10.2136/sssaj1987.03615995005100050015x

Ramankutty, N., Foley, J.A., 1999. Estimating historical changes in global land cover: Croplands from 1700 to 1992. Glob. Biogeochem. Cycles 13, 997–1027. https://doi.org/10.1029/1999GB900046

Sanderson, M.G., 1996. Biomass of termites and their emissions of methane and carbon dioxide: A global database. Glob. Biogeochem. Cycles 10, 543–557. https://doi.org/10.1029/96GB01893

Sobol, I.M., 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math. Comput. Simul., The Second IMACS Seminar on Monte Carlo Methods 55, 271–280. https://doi.org/10.1016/S0378-4754(00)00270-6

Steve, R., Zhao, M., 2021. MODIS/Terra Gross Primary Productivity Gap-Filled 8-Day L4 Global 500m SIN Grid V061." NASA Land Processes Distributed Active Archive Center.

Wang, Y.P., Law, R.M., Pak, B., 2010. A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere. Biogeosciences 7, 2261–2282. https://doi.org/10.5194/bg-7-2261-2010

Zell, J., Kändler, G., Hanewinkel, M., 2009. Predicting constant decay rates of coarse woody debris—A meta-analysis approach with a mixed model. Ecol. Model. 220, 904–912. https://doi.org/10.1016/j.ecolmodel.2009.01.020